HYPERSPECTRAL REFLECTANCE IMAGING FOR DETECTING A FOODBORNE PATHOGEN: *CAMPYLOBACTER*

S. C. Yoon, K. C. Lawrence, G. R. Siragusa, J. E. Line, B. Park, P. W. Feldner

ABSTRACT. *This article is concerned with the development of a hyperspectral reflectance imaging technique for detecting and identifying one of the most common foodborne pathogens, Campylobacter. Direct plating using agars is an effective tool for laboratory tests and analyses of microorganisms. The morphology (size, growth pattern, color, etc.) of colonies grown on agar plates has been widely used to tentatively differentiate organisms. However, it is sometimes difficult to differentiate target organisms like Campylobacters from other contaminants grown together on the same agar plates. A hyperspectral reflectance imaging system operating at the visible and near-infrared (VNIR) spectral region from 400 nm to 900 nm was set up to measure spectral signatures of 17 different Campylobacter and non-Campylobacter subspecies. Protocols for culturing, imaging samples and for calibrating measured data were developed. The VNIR spectral library of all 17 organisms commonly encountered in poultry was established from calibrated hyperspectral reflectance images. A pattern classification algorithm was developed to locate and identify 48 h cultures of Campylobacter and non-Campylobacter contaminants on background agars (blood agar and Campy-Cefex) with over 99% accuracy. The Bhattacharyya distance, a statistical separability measure, was used to predict the performance of the pattern classification algorithm at a few wavelength bands chosen by the principal component analysis (PCA) band weightings. This research has a potential to be expanded to detect other pathogens grown on agar media.*

Keywords. *Agar, Blood agar, Campy-Cefex, Campy-Line agar,* Campylobacter*, Cefex, Contaminant, Food safety, Hyperspectral imaging, Non-*Campylobacter*, Pathogen detection.*

The most commonly recognized foodborne illness is caused by the bacteria *Campylobacter*, *Salmonella*, *E. coli* O157:H7, and Norovirus. The presence of any of these pathogens in food poses a potential hazard to human health. *Campylobacter* is one of the most common causes of bacterial diarrhea illness worldwide. Among 18 different species, *Campylobacter jejuni* is the most commonly isolated species from reported clinical infections, followed by *Campylobacter coli* (USDA-FSIS, 2005). *Campylobacter* in poultry is of particular concern because of the high levels of *Campylobacter* present on most retail chickens, although cattle, unpasteurized milk, and water have also been associated with *Campylobacter* infection. Detection and identification of *Campylobacters* from contaminated food samples involves time-consuming or complicated laboratory tests such as direct plating on agars, immunological techniques using antibody/antigen interactions, or molecular methods using polymerase chain reaction (PCR) or nucleic acid (Stern et al., 2001a). Despite the fact that it is time-consuming, direct plating of samples onto agars has been an effective technique for isolation and enumeration of *Campylobacters* from a variety of sample types; however, distinguishing *Campylobacters* from the non-*Campylobacter* contaminants that frequently grow on many existing agars is difficult (Line, 2001). Recently, use of selective antibiotics and differential chemical dyes has been studied to suppress growth of non-*Campylobacter* contaminants while allowing *Campylobacter* growth (Line, 2001).

In direct plating, growth of bacteria is often characterized by colony morphology. Colonies of different bacteria can vary in size, shape, color, opacity, contrast, shine, etc. The colony morphology can be a good measure for tentatively identifying the bacteria. For example, *Campylobacter* colonies on blood-supplemented or charcoal-based agar media tend to be smooth, convex, and shiny with a distinct edge, or flat, translucent, shiny, and spreading with an irregular edge, as well as colorless to grayish or light cream in color (Stern et al., 2001a). Growth may be confluent without distinct colonies. In addition to the agar type, environmental conditions such as moisture, oxygen level, and temperature also affect the colony morphology. Typically, *C. jejuni* growth on excessively moist agar media swarms, which may be useful to characterize the growth but difficult to isolate individual colonies (Line, 2001). Dry media sometimes promote the growth of only pinpoint (i.e., non-motile) colonies (Line, 2001). Yet, definite identification of the organisms at the species and subspecies levels may require application of immunologically based latex agglutination assays or specific genetic tests like PCR (Stern et al., 2001a).

In addition to visual inspection of the colony morphology, spectral signatures of *Campylobacter* and non-*Campylo bacter*

Submitted for review in December 2007 as manuscript number IET 7317; approved for publication by the Information & Electrical Technologies Division of ASABE in March 2009.

Mention of any company or trade name is for description only and does not imply an endorsement by the USDA.

The authors are **Seung Chul Yoon, ASABE Member Engineer,** Research Electronics Engineer, and **Kurt C. Lawrence,** Research Agricultural Engineer, USDA-ARS Richard B. Russell Research Center, Athens, Georgia; **Gregory R. Siragusa,** Director of Research Food Safety and Quality, Agtech Products, Inc., Waukesha, Wisconsin; **John Eric Line,** Research Food Technologist, **Bosoon Park, ASABE Member Engineer,** Research Agricultural Engineer, and **Peggy W. Feldner,** Food Technologist, USDA-ARS Richard B. Russell Research Center, Athens, Georgia. **Corresponding author:** Seung Chul Yoon, USDA-ARS Richard B. Russell Research Center, Athens, GA 30604-5677; phone: 706-546-3205; fax: 706-546-3607; e-mail: SeungChul.Yoon@ars.usda. gov.

Table 1. Microorganisms used in this study.

Lab Designation	Genus and Species	Strain Designation	Source
1	*Campylobacter jejuni*	ATCC[a] 49943	ATCC
2	*Campylobacter coli*	ATCC 49941	ATCC
3	*Campylobacter lari*	ATCC 43675	Human feces
4	*Campylobacter jejuni*	CT-epi #5	Poultry
5	*Campylobacter jejuni*	CT-epi #65	Poultry
6	*Campylobacter jejuni*	Ty-16C	Poultry carcass rinse
7	*Campylobacter jejuni*	Ty-78	Poultry carcass rinse
8	*Campylobacter coli*	CT-epi #8	Poultry
9	*Campylobacter coli*	CT-epi #18	Poultry
10	*Campylobacter coli*	Ty-19C	Poultry carcass rinse
11	*Campylobacter coli*	Ty-66	Poultry carcass rinse
12	*Sphingomonas paucimobilis*	Contam. 1	Poultry
13	*Acinetobacter baumannii*	Contam. 2	Poultry
14	*Brevundimonas diminuta*	Contam. 3	Poultry
15	*Ochrobacterium* sp.	Contam. 4	Poultry
16	*Flavobacterium odoratum*	Contam. 5	Poultry
17	*Acinetobacter baumannii*	Contam. 6	Poultry

[a] ATCC = American Type Culture Collection, Manassas, Virginia.

colonies may provide vital information about the biochemical compositions of the organisms and thus facilitate the detection and identification of the organisms on agar media. Fourier transform infrared (FT-IR) spectroscopy, which has been applied to solve various microorganism detection/identification problems, was studied to distinguish *Campylobacters* (*C. jejuni* and *C. coli*) at the species level (Mouwen et al., 2005). However, FT-IR spectroscopy is not readily applicable to imaging of Petri dishes. Typically, FT-IR spectroscopy needs an infrared transparent window to mount a biological sample (Naumann, 2000). More readily accessible imaging techniques using a hyperspectral CCD camera are needed to detect and identify *Campylobacter* and non-*Campylobacter* organisms. As a relevant study, a VNIR hyperspectral imaging technique was studied for fungi detection (Yao et al., 2005).

In this article, development of a non-destructive, non-contact hyperspectral imaging technique is discussed to facilitate detection and identification of *Campylobacter* and non-*Campylobacter* colonies grown on Petri dishes of three different agars (blood agar, Campy-Cefex, Campy-Line agar). This research may contribute to build a spectral library of *Campylobacter* and non-*Campylobacter* organisms. The specific objectives of the article were: (1) to develop a protocol for sample preparation and imaging Petri dishes over the VNIR spectral region, (2) to develop classification algorithms that can discriminate *Campylobacters* and non-*Campylobacters* commonly encountered in practice, and (3) to develop an optimal band selection method that is more closely related to the classification accuracy.

Materials and Methods

Microorganism Culture and Preparation of Colony Spot Plates

A collection of *Campylobacter* species and frequently-encountered non-*Campylobacter* species from *Campylobacter* semi-selective agar media were maintained at the USDA-ARS Poultry Microbiological Safety Research Unit's culture collection in Athens, Georgia (table 1). With the exception of the American Type Culture Collection (ATCC) strains, all other bacterial cultures were isolated from poultry samples consisting of either whole-carcass rinses or fecal/cecal specimens from conventionally reared broiler chickens or processing plants (Line, 2001; Siragusa et al., 2004; Stern et al., 2001b). Strains were propagated in 16 mm × 125 mm glass screw-capped culture tubes in 9 mL of *Campylobacter* enrichment broth (Accumedia Catalog No. 7526, Neogen, East Lansing, Mich.) with selective supplements as per the manufacturer's instructions. Tubes were incubated at 42°C for up to 72 h in a Campy-gas atmosphere (85% N_2, 10% CO_2, 5% O_2; Airco Gas, Norcross, Ga.) created by placing the tubes or plates within a re-sealable plastic bag (Reynolds Foodservice Item No. RS 1011, Qwick Seal storage bag, Reynolds Metals Co., Richmond, Va.), gas flushing three times, and refilling with the gas mixture for dense liquid growth. Following the initial liquid culturing step, 5 µL spots (10 µL spots at the early stage of the project) were inoculated to the surface of the respective agar plate and incubated in Campy-gas atmosphere as described at 42°C for a total of 48 h. Agars used were blood agar (5% sheep's blood agar, Remel, Inc., Lenexa, Kans.), Campy-Cefex (Stern et al., 2001b), and Campy-Line agar (Line 2001). Campy-Line agar (CLA) and Campy-Cefex (Cefex) were prepared in-house in standard 100 mm × 15 mm Petri dishes (Siragusa et al., 2004). Cefex agar is one of commonly used agar media to isolate *Campylobacters* from food sources. Cefex is more selective than blood-supplemented agars. CLA is more selective than Cefex and contains a chemical dye, triphenyltetrazolium chloride (TTC), that dyes *Campylobacter* colonies with a deep red or magenta color (Line, 2001).

To avoid confluent growth and cross-contamination, agar plates were inoculated with spots at known and well-spaced locations on the agar surface. Figure 1 shows a schematic of the spot locations and types. The 17 spots were inoculated on two separate plates with nine and eight spots, respectively. The "plate A" plates contained six spots of *Campylobacter* subspecies (Nos. 1 through 6) and three spots of non-*Campylobacters* (Nos. 12 through 14). The "plate B" plates contained five spots of *Campylobacter* subspecies (Nos. 7 through 11) and three spots of non-*Campylobacters* (Nos. 15 through 17).

As implied above, one experiment took place over a five-day period from the sample preparation step (72 h + 48 h) until an

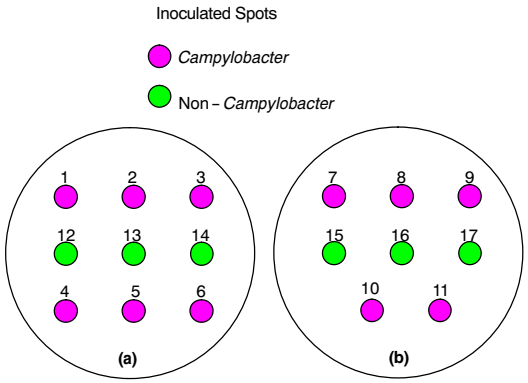


Figure 1. Schematic of spot locations and types: (a) plate A and (b) plate B types. Refer to table 1 for detailed information about the organisms.

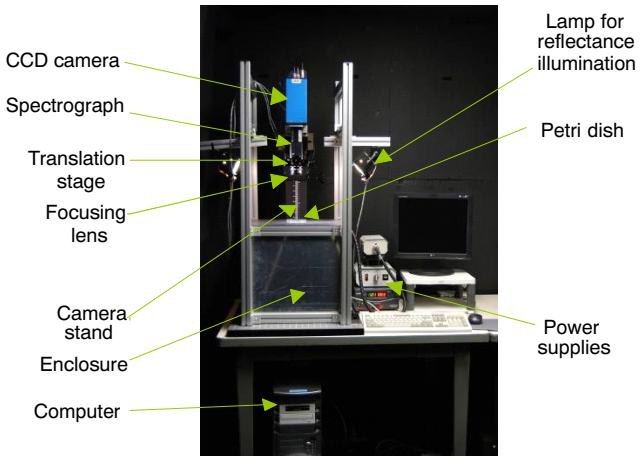


Figure 2. VNIR hyperspectral imaging system. The frontal focusing lens was mounted to a translation stage, which was fixed to the spectrograph. The translation stage was moved to get a hyperspectral image.

imaging step (1 to 5 h). At the end of the second 48 h incubation process, plates were taken out of the incubator and stored at room temperature for 1 to 5 h before being imaged. Eight experiments were carried out from April until July of 2007. For all but two experiments (the first and second ones), experimental protocols for the sample preparation and the imaging remained same. Because we did not know in advance what would be the right amount of inoculation for the imaging study, we prepared one set of 10 µL spot plates and two sets of both 5 µL and 10 µL spot plates for the first and second experiment, respectively. By comparison with 10 µL spots, 5 µL spots were suitable for this study. Hence, we prepared 5 µL spot plates for the last six experiments. In other words, nine sets of spot plates (two of 10 L and seven of 5 µL spot plates) were prepared for a total of eight experiments. The spot plates were prepared in duplicate per agar type and spot size for each experiment (four plates per agar type and spot size). A total of 108 spot plates (for all of three agar types) were prepared for the hyperspectral imaging during the period of four months. The purity of the organisms was checked by visual inspection of grown colonies. Separate plates inoculated by a streaking technique were also prepared to check the organism purity.

Hyperspectral Imaging System

A visible and near-infrared (VNIR) hyperspectral imaging system (fig. 2), similar to one reported earlier (Lawrence et al., 2003), was used for collecting spectral and spatial images. The imaging system consisted of a hyperspectral imaging camera, a copy stand to attach the camera, a computer to control the camera and acquire images, an enclosure to block unwanted light, halogen lamps, and a Petri dish holder. The hyperspectral imaging camera (ITD, Stennis Space Center, MS) consisted of a C-mount focusing lens (XNP 1.4/17-0303, Schneider Optics, Hauppage, N.Y.), a spectrograph (ImSpector V10E, Specim, Oulu, Finland), a 12-bit CCD sensor (SensiCam QE SVGA, Cooke Corp., Auburn Hills, Mich.), and associated scanning hardware. The hyperspectral imaging camera was designed so that the target and the camera remained stationary while the lens assembly moved (Mao, 2000). The spectrograph had a prism-grating-prism design providing a nominal spectral range between 400 and 1000 nm with a nominal spectral resolution of 2.8 nm and bandpass of 2.95 nm, and it was connected to the 17 mm (2/3 in.) silicon-based CCD detector with a 1280 × 1024 pixel resolution. A translation stage (STGA-10, Newmark Systems, Mission Viejo, Cal.) was attached to the slit end of the spectrograph. The stage platform was threaded for a C-mount focusing lens. The translation stage was moved by a motion controller (NCS-1S, Newmark Systems, Mission Viejo, Cal.). Thus, the motorized translation stage moved the lens assembly so that successive lines of the Petri dish were scanned while the Petri dish itself remained stationary.

The hyperspectral imaging system captured two-dimensional (2-D) spectral images of the scene in the field of view line-by-line by moving the frontal optic lens via the translation stage (Lawrence et al., 2003). The translation stage moved continuously while a hyperspectral image was acquired. Operation parameters (scan speed and travel length) of the translation stage were affected by camera exposure time, and the optimal operation parameters were pre-determined by ITD and automatically configured within HyperVisual software (ITD, Stennis Space Center, Miss.) once a camera exposure time was set. The captured spectral images at the x and z coordinates were synthesized to a three-dimensional (3-D) data cube of the x, y, and z coordinates via HyperVisual on the fly, where x and y are the spatial coordinates and z is the wavelength coordinate. For the reflectance illumination, two 50 W halogen MR16 lamps (Solux, EiKo, Shawnee, Kans.) were positioned at the locations obliquely pointing down the Petri dish from the left and right sides. The top of the enclosure was covered with a metal plate having a hole at its center to hold a Petri dish. A white acrylic diffuser was placed right beneath the cover plate, and a Petri dish was put on top of the white acrylic diffuser. The data analysis and algorithm development were performed using IDL 6.3/ENVI 4.3(ITT Visual Information Solutions, Boulder, Colo.) and MATLAB R2006b (The Mathworks, Natick, Mass.).

Measurement and Calibration

Two 4700 K broadband light beams laterally illuminated a Petri dish at about 45° with respect to the vertical axis. Lateral illumination was chosen rather than a 90° incident angle because the lateral illumination reduced glare effects better. The actual beam angles and locations of the lamps were adjusted to avoid pixel saturation on the CCD sensor and to minimize glare reflections (glints) from the agar surface and colonies. The distance from the Petri dish to the camera lens was 22 cm. The distance from the Petri dish to a lamp for front lighting was 37 cm. Two-dimensional spectral images (i.e., line-scan images) were captured by 2 (spatial) × 4 (spectral) software binning and 90 ms exposure time. The F-number of the lens was set to 4. The

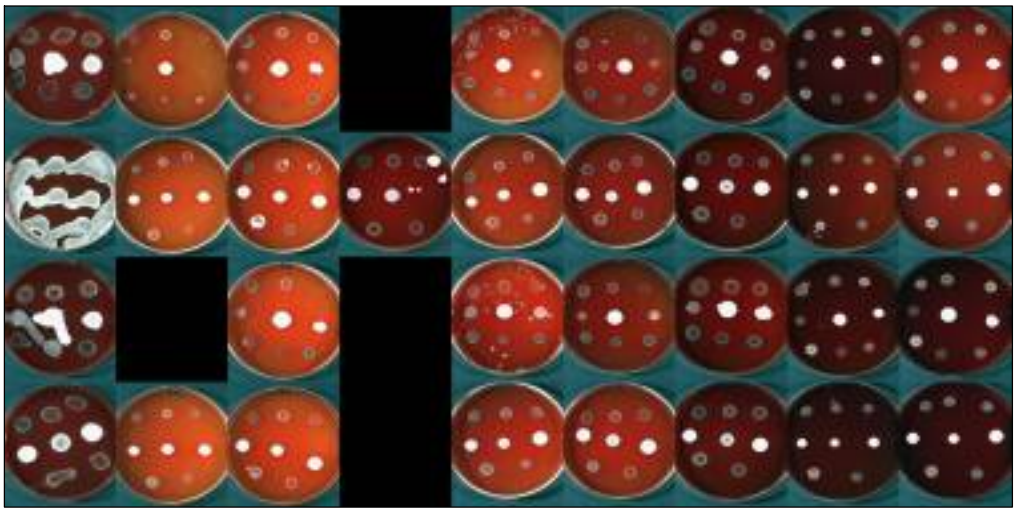


Figure 3. Hyperspectral image mosaic (Cefex cultures). Color-composite images are displayed.

field of view (FOV) was 115 mm (W) × 93 mm (H). The pixel resolution of a line-scan image was 640 (spatial) × 256 (spectral). A total of 475 lines were scanned. Hence, the dimension of the 3-D image data cube was 640 (W) × 475 (H) × 256 (wavelength).

Acquired images were pre-processed to reduce the amount of data and to suppress spurious spectral noise. All band images were cropped down to 421 (W) × 475 (H). The 193 spectral bands in the range of 400 to 900 nm were kept because the spectra outside the range were not useful enough to provide viable spectral information due to both the low diffraction efficiency of the spectrograph and the low quantum efficiency of the CCD detector. Typical CCD sensors are prone to random noise due to various reasons, including dark current, readout noise, shot noise, etc. The Savitzky-Golay smoothing filter (window size: 25; order of moment: 4) was applied to each pixel independently to reduce the spectral random noise (Press et al., 2002).

The hyperspectral camera was spectrally calibrated (Lawrence et al, 2003). The average distance between adjacent calibrated wavelengths, defined as the measured spectral resolution, was 2.5953 nm. The spectral binning done by software increased the measured spectral resolution by a factor of 2. New wavelengths of the binned data were derived from the original calibrated wavelengths by taking an arithmetic mean over the disjoint sets of wavelengths. For instance, the spectral binning size of 4 resulted in a spectral resolution of about 10 nm. Hence, we used integer values for wavelengths throughout this article for brevity (e.g., 501 nm instead of using 500.8901 nm) because the fractional representation of wavelengths was not critical. Intensity calibration was performed with a 75% reflectance Spectralon target (13 × 13 cm, SRT-75-050, Labsphere, North Sutton, N.H.) and a reflectance calibration model (Lawrence et al., 2003). The percent reflectance value R at each pixel (x,y) of the zth wavelength band was obtained by the following calibration model:

$$R(x,y,z) = \frac{I_m(x,y,z) - I_d(x,y,z)}{I_r(x,y,z) - I_d(x,y,z)} \cdot 75 \qquad (1)$$

where I_m is a measured raw value, I_r is a reference value on the surface of the 75% reflectance panel, and I_d is a dark current.

Data Analysis

Calibrated images were arranged into a single image (mosaic) according to their measured dates (fig. 3). An image mosaic approach was adopted to facilitate data analysis and algorithm development because a mosaic can be treated as a single hyperspectral image. In the mosaic, images measured during the same date were vertically stacked. The stacked images were then added to the mosaic in chronological order (latest right). The "plate A" type plates were arranged in the first row. The "plate B" type plates were arranged in the second row. The third and fourth rows were replicates of the top two rows, respectively. Figure 3 shows an image mosaic of Cefex cultures. Note that, throughout this article, the Cefex cultures will be used as a representative of the three agar types unless otherwise stated.

A binary mask was made in order to suppress the background noise outside and around the rim of each Petri dish (fig. 4). These binary masks were served as areas valid for testing classification algorithms (i.e., cross-validation of classification algorithms) and for facilitating other tasks, including image processing and analysis. In addition, ground-truth regions-of-interest (ROIs) for all 17 organisms plus agar media were prepared in such a way that only pure organisms could be

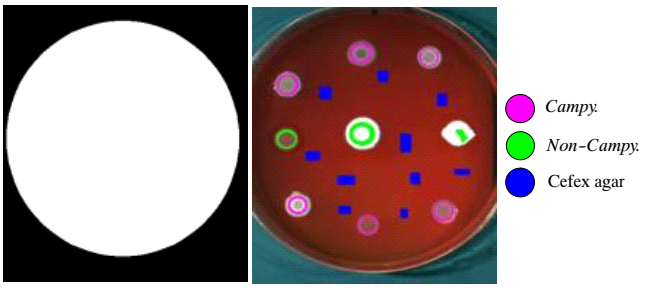


Figure 4. Image mask for cross-validation (left) and ground-truth ROIs (right) with training classes.

selected in the ROIs. Glints and rim shadows were excluded in the ROIs. Glints were observed around colony edges along the direction of the lateral illumination. If possible, pixels that might have contained mixed spectra of agar media and organisms were also excluded. Mixed pixels were typically observed at the center of a spot with a translucent colony surface and at outside boundaries of a spot. As an example, figure 4 shows a binary mask for cross-validation and ground-truth ROIs of three classes: *Campylobacter*, non-*Campylobacter*, and Cefex agar. The selected ground-truth ROIs were modified by the aforementioned manual method, if necessary, after heuristically examining the shapes of the mean, minimum, maximum, and standard deviation reflectance spectra of each ROI. The final ground-truth ROIs were used for extracting statistics and training classifiers.

Band Selection and Statistical separability

In practice, due to the usual limitation on the number of training samples, more spectral features do not necessarily lead to better classification results. This phenomenon is called the Hughes effect (Landgrebe, 2003). In addition, spectral features in hyperspectral images contain a lot of redundant information because absorption features in adjacent spectral bands are highly correlated. Hence, there is a need to find optimal, smaller subsets of spectral features to use in classification. In this study, principal component analysis (PCA) was used to determine which spectral bands contribute to most of the data variability (Landgrebe, 2003). A statistical separability test was performed on the spectral bands chosen by the PCA. A PCA model consists of principal components (or often called loadings or eigenvectors for PCA), PCA-bands (or scores), and eigenvalues. Principal components contain information about directions in which the original data have the most variability. Mathematically, principal components correspond with the eigenvectors of the covariance matrix of the mean-centered data, and they are uncorrelated and orthogonal. PCA-bands (i.e., scores) are obtained by projecting the data onto the principal components. Mathematically, a score is the scalar product of a data vector and a principal component. Thus, elements of a principal component are weighting factors individually contributing to scores. The principal components were arranged in descending order according to the associated eigenvalues. To determine the contribution of each of all n bands to the scores (a PCA image band in our case), we simply computed squares of principal components and normalized to the sum of 1 by the following equation:

$$W_k(i) = P_k(i)^2 / \sum_{i=1}^{n} P_k(i)^2,\ i \in \{i, \Lambda, n\} \qquad (2)$$

where $W_k(i)$ is a weighting factor of the ith element of the kth principal component, $P_k(i)$. We call all of $W_k(i)$ the PCA-band weightings. The PCA-band weightings of the first and second principal components were examined.

As another band selection method, the statistics of all spectra observed in each type of organism were computed. Mean, minimum-maximum, and standard deviation spectra were examined to find statistical separability at particular spectral bands. Since about 68% of the normally distributed population is within one standard deviation of the mean (95% for two standard deviations of the mean), we can roughly estimate class separability by examining minimum-maximum and standard deviation intervals from the class mean. The ground-truth ROI mask was used for this test. The candidates of the bands having large separability may be chosen by this method. Note that the PCA-based band selection does not explicitly measure class separability that is more directly correlated to classification accuracy because the PCA only tells the statistical variability among spectral bands (not separability among classes). On the other hand, the band selection method using the intervals from the mean is somewhat heuristic. Because of these limitations, we studied a statistical separability measure predicting classification accuracy at a few bands chosen by the aforementioned band selection processes.

Among many different statistical separability metrics, the Bhattacharyya distance measures the dissimilarity of two probability distributions, and it has been used in classification as a measure for the separability of classes. The Bhattacharyya distance is linearly related to the probability of correct classification (Comaniciu et al., 2000; Landgrebe, 2003). Assuming that $\hat{p}_i(x)$ is an estimate of the discrete probability density function of the ith class at x, where x is an intensity value, the Bhattacharyya distance between the ith class and jth class can be estimated by:

$$\hat{d}(i, j) = 1 - \sum_{x} \sqrt{\hat{p}_i(x)\hat{p}_j(x)} \qquad (3)$$

We used this statistical measure (eq. 3) for measuring the statistical separability of a pair of two classes among *Campylobacter*, non-*Campylobacter* contaminant, and agar-media classes in several bands chosen from the PCA-band weighting analysis mentioned above. The derivation of the Bhattacharyya distance needs the estimation of the probability density functions, for which we employed the histogram formulation (Comaniciu et al., 2000). The implementation of equation 3 based on a histogram method is fast and well suited for the task of measuring class separability because the Bhattacharyya distance measure is nearly optimal and has a theoretical foundation (Comaniciu et al., 2000). The Bhattacharyya distance varied from 0 (minimum) to 1 (maximum).

Classification

Our primary concern was to find which agar type and which image classification method would be the best to distinguish between *Campylobacters* and non-*Campylobacters* grown on agar plates. Given the measured data, a preliminary study found the Cefex and the blood agar to be the best possible agar media to separate the two groups of organisms. Hence, this article reports the findings primarily based on Cefex and blood agar. For both agar types, three classes (*Campy.*, non-*Campy.*, and agar) were initially designed for classification. For blood agar cultures, a simple threshold-based classifier was developed to classify pixels of a single band into the three classes and glints. The selection of the band and thresholds was necessary. On the other hand, a similar threshold-based classifier was applied to Cefex cultures, but we found that there was a hard-to-distinguish organism (a non-*Campylobacter* contaminant; sample No. 12). Thus, we developed a detection algorithm for better classification of the Cefex cultures by compensating this limitation of the single-band thresholding algorithm. The developed algorithm was a two-step hybrid method that performed (1) a threshold-based classifier on a single band for classifying the three classes and the glints and (2) a two-class minimum distance classifier using the Mahalanobis distance

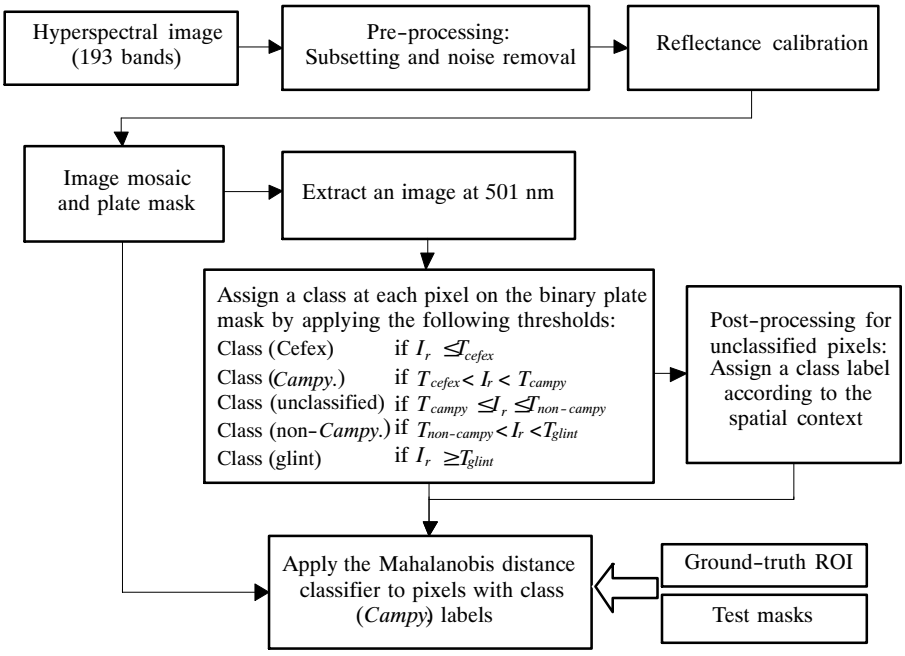


Figure 5. Overview of the detection algorithm for Cefex cultures: I_r is a reflectance value at a pixel.

measure, which used all 193 spectral bands. All pixels detected as the *Campy.* class during the single-band thresholding step were classified to the closest class, *Campylobacter* or sample No. 12. A preliminary study to determine the input size of the Mahalanobis distance classifier for classifying the *Campy.* class and sample No. 12 revealed that the classification performance of all 193 bands was better than the performance of three bands determined by the PCA-band weighting analysis.

Figure 5 summarizes the detection algorithm for Cefex cultures. First, each hyperspectral image was re-sized and spectrally de-noised by the methods mentioned earlier. After the reflectance calibration, all calibrated images were stitched into a mosaic similar to one shown in figure 3. Another mosaic of binary plate masks (fig. 4) was manually generated for confining the classification operation inside the Petri dish areas only. Then a threshold-based classifier was applied to the masked image mosaic at 501 nm to separate the three classes (*Campy.*, non-*Campy.*, and Cefex) and glints. At each pixel, the classifier worked as follows: (1) Cefex if I_r (a reflectance value) $\leq T_{cefex}$, (2) *Campylobacter* if $T_{cefex} < I_r < T_{campy}$, (3) unclassified if $T_{campy} \leq I_r \leq T_{non\text{-}campy}$, (4) non-*Campylobacter* if $T_{non\text{-}campy} < I_r < T_{glint}$, and (5) glint if $I_r \geq T_{glint}$. The glint pixels were assigned to be unclassified. The threshold values for classifying the pixels in the 501 nm band image were determined by trial and error. Class means and histograms of the reflectance values were examined to get initial guesses. The final threshold values were found to be close to the midpoints of class-mean intervals. To further classify unclassified pixels on Cefex cultures, spatial contexts of the unclassified pixels were examined. If a group (blob) of spatially connected unclassified pixels was surrounded or connected by one particular type of class label, then that particular class label was assigned to the group of pixels. If there was more than one type of class labels, then the unclassified group of pixels remained unclassified. After this post-processing, the Mahalanobis distance classifier was applied to only the predicted *Campylobacter* class in order to separate a particular class (sample No. 12) from the predicted *Campylobacter* class. A leave-one-out cross-validation method was adopted to evaluate the performance of the Mahalanobis distance classifier (Webb, 2002). For this cross-validation, all but one day's data were used for training the classifier, and the rest (one day's data) were left out and used for testing (validating) the classifier. Note that this cross-validation method was applied to only the predicted *Campylobacter* pixels obtained by the single-band thresholding algorithm for Cefex cultures shown in figure 5.

Results and Discussion

Colony Size and Patterns

Figure 6 shows an example of digital camera photos of Petri dishes of colony cultures grown on blood agar, Cefex, and CLA plates. Because it was difficult to exactly control the amount (5 μL/10 μL) of each spot for every experiment, inoculation spots varied in size. In part, this led to colonies grown to different sizes. The size of colonies grown on blood agar plates ranged from 9 to 15 mm in diameter. Cefex agar plates produced colonies of 6 to 12 mm in diameter. Colonies grown on CLA plates had diameters from 7 to 11 mm. Some spots did not grow as colonies at all. Others grew too much and spread into neighboring spots, causing cross-contamination. Some inoculation spots suffered from cross-contamination due to the impurity of samples in stock. In this case, different types of colonies were observed from one spot. Every effort was made to maintain the pipeline of pure organisms and to produce pure cultures. Apart from the impurity factor added during the sample preparation, the type of an agar plate had the most significant influence on colony growth pattern and colony color (spectra in our case). Colonies on blood agar plates tended to spread more compared to colonies on Cefex and CLA plates. On the other hand, Cefex and CLA plates produced less spread colonies with round and raised shapes. Sometime, a ring-type growth pattern was observed because the colonies were growing outward. Figure 3 shows a color-composite image mosaic of

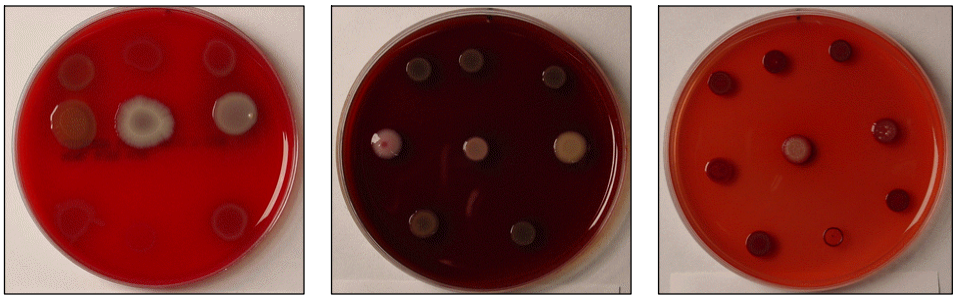


Figure 6. Grown culture example: blood agar, Cefex, and CLA (from left).

Cefex cultures. The black boxes in the mosaic mean missing plates (not imaged) because of confluent growth or no growth. One example of confluent growth is shown in the top second and third plates in the first column of the mosaic (fig. 3). The first and third columns contain plate images with 10 μL inoculation spots. The second (5 μL) and third (10 μL) column images were taken on the same date. The blood agar and CLA cultures, not shown in this article due to limited space, also suffered from the same problems: cross-contamination, no-growth, or confluent growth.

Spectral Analysis

Mean and standard deviation reflectance spectra were obtained from the ground-truth ROI masks of all 17 organisms grown on the three types of agars over the 400-900 nm range, as shown in figure 7. Blood agar (BA) and Cefex showed similar mean spectral characteristics. In both cases, all of the 11 *Campylobacter* cultures showed low reflectance responses (~5%) over the visible spectra below 650 nm (figs. 7a and 7c). The near-infrared spectra showed larger reflectance (~35%). On the other hand, non-*Campylobacter* cultures had a strong absorption feature at around the green band (550 nm) and much larger reflectance features over the visible spectral range than the spectra of the *Campylobacter* cultures. Beyond 700 nm, non-*Campylobacters* (except sample No. 12, *Sphingomonas paucimobilis*, in the Cefex case) had larger reflectance than *Campylobacters* (figs. 7a and 7c). The mean spectrum of the sample No. 12 contaminant grown on Cefex agar was very similar to that of *Campylobacters*. The mean spectrum of the non-*Campylobacter* cultures grown on CLA was different from the other two agars in the sense that it was much lower (less than 10%) below 560 nm, but it was much higher (larger than 40%) in the rage from 700 nm to 850 nm (fig. 7e).

From the standard deviation plots in figures 7b, 7d, and 7f, it was obvious that the three classes (*Campy.*, non-*Campy.*, and agar) of blood agar cultures and Cefex cultures were visually well separable at around 500 nm, whereas the classes of CLA cultures were not separable. From figures 7b and 7d, the separability of the data around 500 nm was much larger than one standard deviation. Note that sample No. 12 was not readily separable (fig. 7d), and it was treated separately using the detection algorithm. The spectral variability of BA and Cefex cultures increased starting from 550 nm. In addition, in all agar cases, it was observed that reflectance values of all organisms sharply increased beyond 650 nm until 750 nm and then slowly decreased after 800 nm. Overall, the variability of the spectra in the range of 700-900 nm became larger, and the mean spectra over the same range were overlapped. From the above observations, we hypothesized that the target organisms would be statistically better separable in the subrange of 450-530 nm within the 400-900 nm range than the rest. This hypothesis was evaluated by estimating the Bhattacharyya distances at a few bands determined by the PCA-band weight analysis.

Band Selection and Statistical Separability

PCA was applied to the ground-truth reflectance spectra of blood agar, Cefex, and CLA cultures. About 76% and 20% of the variation energy (measured by the magnitude of the first- and second-largest eigenvalues) were captured by the first two PCA-bands (scores) in the blood agar case, and 78% and 16% in the Cefex case. The weight vectors (see eq. 2) contributing to the first PCA-bands are shown in figures 8a and 8b (note that the CLA case is omitted.). Two bands at 503 and 578 nm (the blood agar case) and three bands at 501, 606, and 827 nm (the Cefex case) were observed as local peaks in the band contribution plots, respectively. In the case of CLA cultures, a band at 633 nm and any band between 750 and 900 nm were candidates. A scatter plot of the first two PCA weight vectors provided another good perspective for band selection (figs. 8c and 8d). Similarly, two bands at 503 and 578 nm (the BA case) and three bands at 501, 606, and 827 nm (the Cefex case) were valid candidates. It was interesting to observe that the valleys at 691 nm (fig. 8a) and 685 nm (fig. 8b) on the first PCA-band weighting axis showed the largest weighting values (crests) on the second PCA-band weighting axis (figs. 8c and 8d). Class separability in each of these bands except ones at the valleys was measured by the Bhattacharyya distance.

Table 2 summarizes the Bhattacharyya distances among three classes (*Campy.*, non-*Campy.*, and agar) of BA, Cefex, and CLA cultures. For BA cultures, the band at 503 nm showed the largest separability between *Campylobacter* and non-*Campylobacter*. For Cefex cultures, the band at 501 nm was the best in separating all three classes. In the CLA case, the overall separability was smaller than the other agars. *Campylobacter* and all agars were well separable at 503 nm (BA), 501 nm (Cefex), and 633 nm (CLA). The Bhattacharyya distance between *Campylobacters* and non-*Campylobacters* on Cefex was lower than BA because of sample No. 12. When sample No. 12 was treated as an independent class, the Bhattacharyya distance became 0.950485 from 0.837521 at 501 nm. As a reference, the Bhattacharyya distance between *Campylobacters* and sample No. 12 on Cefex was 0.238246 at 501 nm, which means that they were not statistically separable. Remember that the Bhattacharyya distance is almost linearly related to classification accuracy. Thus, high classification accuracy was expected when a single band at either 501 nm (Cefex) or 503 nm (BA) was used. In the following section, classification accuracy of blood agar first and then Cefex cultures is reported.

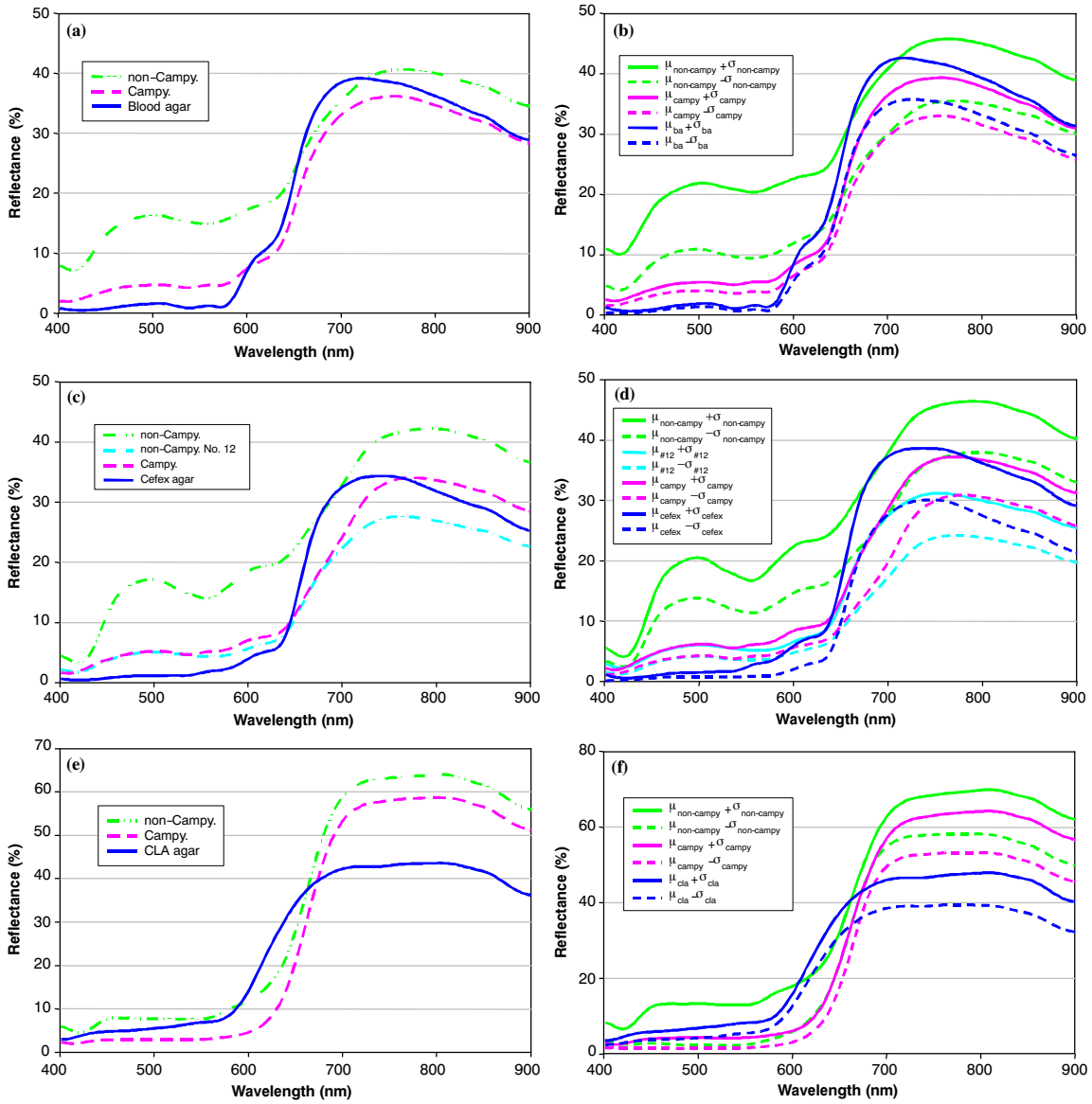
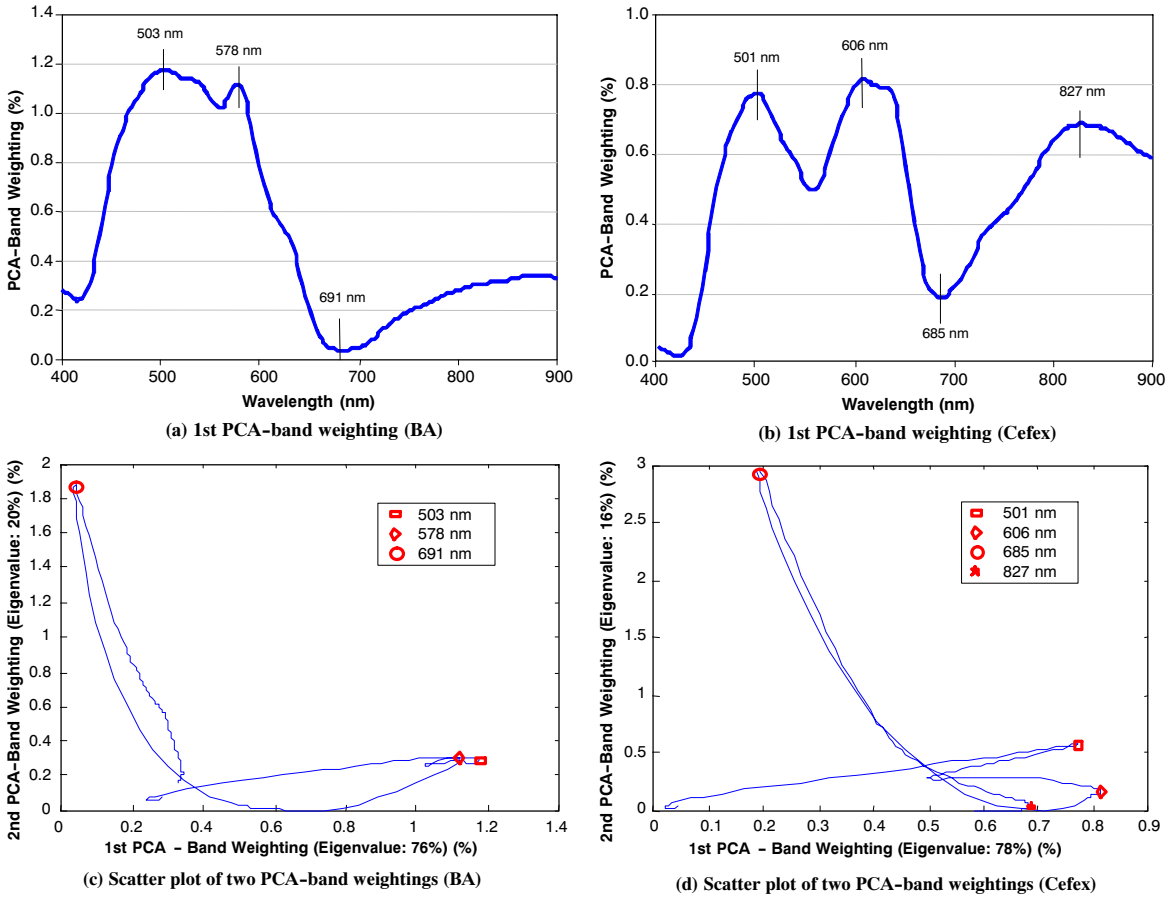


Figure 7. Mean and standard deviation reflectance spectra of (a and b) BA cultures, (c and d) Cefex cultures, and (e and f) CLA cultures. The ROI masks were used to compute the statistics of mean and standard deviations of reflectance spectra. The Cefex had four classes, whereas the others had three classes.

Classification Accuracy: Blood Agar and Cefex

The classification of the 503 nm band image of blood agar (BA) cultures was done by the following rule: (1) blood agar if a reflectance value is below 3, (2) non-*Campylobacter* if the value is greater than 7, and (3) *Campylobacter* otherwise. Table 3 summarizes the classification accuracy. For computing the confusion matrix in table 3a, sample No. 12 (*Sphingomonas paucimobilis*, contaminant 1) was treated as a separate set of ground-truth ROI pixels. When the *Sphingomonas paucimobilis* class was included in the non-*Campylobacter* class, the total number of the ground-truth ROI pixels was 136,370 and the overall classification accuracy was 98.07% (133,740 pixels/136,370 pixels). The Kappa coefficient was 0.9703. The BA ROI pixels (n = 59,017) were classified with 100% accuracy. The *Campylobacter* ROI pixels (n = 32,590) were classified with 98.97% accuracy (32,254 pixels/32,590 pixels). The non-*Campylobacter* ROI pixels (n = 44,763) were classified with 94.88% accuracy (42,469 pixels/44,763 pixels), whereas sample No. 12 alone (n = 4,200) was classified with 52.90% accuracy (2,222 pixels/4,200 pixels). Almost half (47.10%) of the sample No. 12 ROI pixels were misclassified as *Campylobacter* class. The commission errors of the *Campylobacter*, non-*Campylobacter*, and BA classes were 6.64%, 0.11%, and 0.49%, respectively (table 3b). The omission errors of the *Campylobacter*, non-*Campylobacter*, and BA classes were 1.03%, 5.12%, and 0%, respectively (table 3b). Sample No. 12 was mainly responsible for the relatively large commission error and omission error rates (6.64% and 5.12%) of the *Campylobacter* and non-*Campylobacter* classes, respectively. A secondary factor that affected the errors was a spectral mixing phenomenon around the inner and outer rims of a colony in which the contribution of the background agar to the spectral response was not trivial. When the sample No. 12 ROI pixels were excluded from the non-*Campylobacter* ground-truth ROI pixels, the total number of ground-truth ROI pixels became 132,170 and the overall classification accuracy increased to 99.51% (131,518 pixels /132,170 pixels). The



(a) 1st PCA-band weighting (BA)

(b) 1st PCA-band weighting (Cefex)

(c) Scatter plot of two PCA-band weightings (BA)

(d) Scatter plot of two PCA-band weightings (Cefex)

Figure 8. PCA-band contribution weightings of blood agar and Cefex cultures. Wavelengths were rounded up to the nearest integer values.

Table 2. Bhattacharyya distance.

	Blood agar		Cefex			CLA	
	503 nm	578 nm	501 nm	606 nm	827 nm	633 nm	816 nm
Campy. vs. non-*Campy.*	0.927	0.912	0.838	0.831	0.613	0.670	0.362
Campy. vs. agar	0.999	1.000	0.999	0.591	0.285	0.994	0.848
Non-*Campy.* vs. agar	1.000	0.993	1.000	0.893	0.663	0.741	0.922

Table 3. Classification accuracy (the single-band thresholding algorithm only): blood agar (BA).

(a) Confusion matrix

	Ground Truth (Pixels)				
Class	*Campylobacter*	Non-*Campy.*[a]	No. 12 (non-*Campy.*)	BA	Total
Campylobacter	32,254 (98.97%)	316 (0.78%)	1,978 (47.10%)	0 (0%)	34,548 (25.33%)
Non-*Campy.*	45 (0.14%)	40,247 (99.22%)	2,222 (52.90%)	0 (0%)	42,514 (31.18%)
BA	291 (0.89%)	0 (0%)	0 (0%)	59,017 (100%)	59,308 (43.49%)
Total	32,590 (100%)	40,563 (100%)	4,200 (100%)	59,017 (100%)	136,370 (100%)

(b) Commission and omission errors when sample No. 12 was included in the non-*Campy.* class

Class	Commission (%)	Omission (%)	Commission (pixels)	Omission (pixels)
Campylobacter	6.64	1.03	2,294/34,548	336/32,590
non-*Campy.*	0.11	5.12	45/42,514	2294/44,763
BA	0.49	0.00	291/59,308	0/59,017

(c) Commission and omission errors when sample No. 12 was excluded from the non-*Campy.* class

Class	Commission (%)	Omission (%)	Commission (pixels)	Omission (pixels)
Campylobacter	0.97	1.03	316/32,570	336/32,590
non-*Campy.*	0.11	0.78	45/40,292	316/40,563
BA	0.49	0.00	291/59,308	0/59,017

[a] All non-*Campylobacter* contaminants except sample No. 12.

Kappa coefficient became 0.9924. The BA ROI pixels were classified with 100% accuracy. The *Campylobacter* ROI pixels were classified with 98.97% accuracy. The non-*Campylobacter* ROI pixels without the sample No. 12 ROI pixels were classified with 99.22% accuracy. The commission errors of the *Campylobacter*, non-*Campylobacter*, and BA classes became 0.97%, 0.11%, and 0.49%, respectively (table 3c). The omission errors of the *Campylobacter*, non-*Campylobacter*, and BA classes were 1.03%, 0.78%, and 0%, respectively (table 3c).

Similarly to the single-band thresholding algorithm applied to the BA culture images, the Cefex culture images were classified by the following single-band thresholding algorithm: (1) Cefex if a reflectance value is below 3, (2) non-*Campylobacter* if the value is greater than 10, and (3) *Campylobacter* otherwise. Table 4a summarizes the performance of the single-band thresholding algorithm. The detection accuracy of the *Campylobacter* class (n = 93,224) was 98.99% (92,282 pixels/93,224 pixels) with 7.51% (7,498 pixels/99,780 pixels) of the commission error rate. The detection accuracy of the Cefex class (n = 96,372) was 100% with 0.98% (957 pixels/97,329 pixels) of the commission error rate. The detection accuracy of the non-*Campylobacter* class (n = 75,388) was 90.03% (67,875 pixels/75,388 pixels) without commission errors (0 pixels/67,875 pixels). The factors affecting the omission and commission errors were similar to those of the BA case, where sample No. 12 (n = 5,837) was a major reason for the errors and the spectral mixing problem around the inner and outer rims of a colony was a secondary factor. If the confusion matrix of table 4a is carefully examined, you will notice that the detection accuracy of sample No. 12 was 0% (0 pixels/5,837 pixels). In addition, 99.74% (5,822 pixels/5,837 pixels) and 0.26% (15 pixels of 5,837 pixels) of the sample No. 12 ROI pixels were classified as *Campylobacter* class and Cefex class, respectively. This result strongly suggests that the pixels in the predicted *Campylobacter* class needed to be classified again in order to extract the sample No. 12 class pixels. In comparison, only half (47.10%) of the sample No. 12 ROI pixels needed to be further classified for the BA case. Due to the complete misclassification of the sample No. 12 pixels, an image classification algorithm was developed for the Cefex culture images only, and its results are reported in the following.

The threshold parameters of the developed algorithm for classifying the 501 nm band of Cefex cultures were set as follows: $T_{cefe\ x}$ = 3, $T_{campy.}$ = 7.5, $T_{non\text{-}campy.}$ = 10, and $T_{glint.}$ = 30. A reflectance value between 7.5 and 10 was not classified initially. After post-processing, the unclassified pixels were labeled as described before. The pixels predicted as *Campylobacter* class were further classified by the two-class Mahalanobis distance classifier to separate *Sphingomonas paucimobilis* (sample No. 12) from *Campylobacters*. The resulting classification result was quantitatively evaluated at the ground-truth ROI pixels. The total number of the ground-truth ROI pixels was 264,984: *Campylobacter* (93,224), non-*Campylobacter* (75,388), Cefex (96,372). The classification performance was summarized in a confusion matrix shown in table 4b. The overall classification accuracy was 99.29% (263,104 pixels/264,984 pixels). The Kappa coefficient was 0.9893. The Cefex ROI pixels were classified with 100% accuracy. The *Campylobacter* ROI pixels were classified with 98.62% accuracy. The non-*Campylobacter* ROI pixels were classified with 99.22% accuracy. The commission errors of the *Campylobacter*, non-*Campylobacter*, and Cefex classes were 0.3%, 0.3%, and 1.19%, respectively. The omission errors of the *Campylobacter*, non-*Campylobacter*, and Cefex classes were 1.38%, 0.78%, and 0%, respectively. Only 0.08% of pixels were still unclassified. Most of the unclassified pixels were observed in one spot, where the colony was translucent, small, and cross-contaminated. The spectral similarity of *Sphingomonas paucimobilis* and *Campylobacter* was responsible for most misclassified pixels. Next, the developed classification algorithm was qualitatively evaluated on the entire image space

Table 4. Classification accuracy (the single-band thresholding algorithm and the developed algorithm): Cefex.

(a) Confusion matrix of the single-band thresholding algorithm

	Ground Truth (Pixels)				
Class	*Campylobacter*	non-*Campy.*[a]	No. 12 (non-*Campy.*)	Cefex	Total
Campylobacter	92282 (98.99%)	1676 (2.41%)	5822 (99.74%)	0 (0%)	99780 (37.66%)
non-*Campy.*	0 (0%)	67875 (97.59%)	0 (0%)	0 (0%)	67875 (25.61%)
Cefex	942 (1.01%)	0 (0%)	15 (0.26%)	96372 (100%)	97329 (36.73%)
Total	93224 (100%)	69551 (100%)	5837 (100%)	96372 (100%)	264984 (100%)

(b) Confusion matrix of the developed algorithm

	Ground Truth (Pixels)				
Class	*Campylobacter*	non-*Campy.*[a]	No. 12 (non-*Campy.*)	Cefex	Total
Unclassified	16 (0.02%)	131 (0.19%)	69 (1.18%)	0 (0%)	216 (0.08%)
Campylobacter	91935 (98.62%)	156 (0.22%)	124 (2.12%)	0 (0%)	92215 (34.80%)
non-*Campy.*[b]	0 (0%)	68657 (98.71%)	0 (0%)	0 (0%)	68657 (25.91%)
No. 12	228 (0.24%)	511 (0.73%)	5629 (96.44%)	0 (0%)	6368 (2.40%)
Cefex	1045 (1.12%)	96 (0.14%)	15 (0.26%)	96372 (100%)	97528 (36.81%)
Total	93224 (100%)	69551 (100%)	5837 (100%)	96372 (100%)	264984 (100%)

(c) Commission and omission errors of the developed algorithm

Class	Commission (%)	Omission (%)	Commission (Pixels)	Omission (Pixels)
Campylobacter	0.30	1.38	280/92215	1289/93224
non-*Campy.*	0.30	0.78	228/68657	591/75388
Cefex	1.19	0.00	1156/97528	0/96372

[a] All non-*Campylobacter* contaminants (ground-truth) except the sample No. 12.
[b] Predicted class for all non-*Campylobacter* contaminants except the sample No. 12.

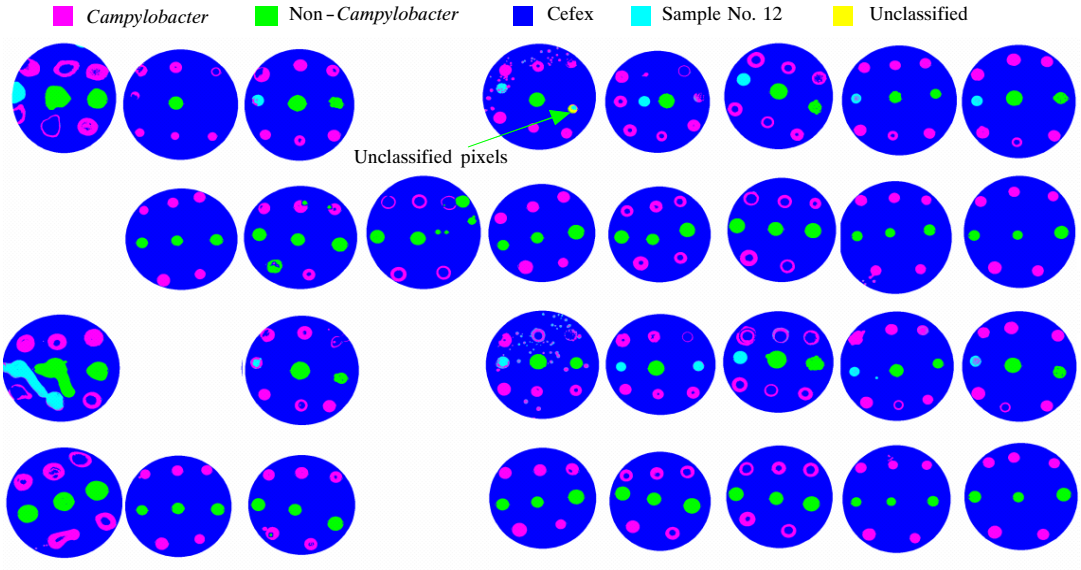


Figure 9. Classification result of Cefex cultures.

after performing leave-one-out cross-validation. The classification images were combined into a mosaic. Figure 9 shows a mosaic of classification images tested by the cross-validation process. At the spot level, all spots except one were correctly classified. The misclassified spot contained three labels: *Campylobacter*, non-*Campylobacter*, and unclassified. The unclassified pixels consisted of the majority of the spot. Other common misclassified pixels were observed either from the *Campylobacter* and *Sphingomonas paucimobilis* classes because of the inefficiency of the Mahalanobis distance classifier using the entire 193 spectral bands.

Conclusions

A VNIR hyperspectral reflectance imaging technique was evaluated to develop an imaging protocol to detect *Campylobacter* species and non-*Campylobacter* species commonly encountered in poultry carcasses in a controlled condition. The 48 h cultures from 5 µL spots were used for the study. The spectral responses of the colonies were examined over the range of 400 nm to 900 nm. Blood agar and Campy-Cefex agar were selected to develop the detection algorithm. Several spectral bands (501, 503, 606, and 827 nm) showed large statistical variability. The Bhattacharyya distance measure was employed to select optimal bands showing the largest statistical separability among target classes. The band at around 500 nm showed the largest separability for both blood agar and Cefex cultures because most of the pixels were correctly classified by segmenting the single band image with thresholds. However, in the case of Cefex cultures, one particular non-*Campylobacter* species had spectral responses similar to *Campylobacter* species. To classify the two species, a statistical classifier (Mahalanobis distance classifier) was evaluated over the entire spectral band with a leave-one-out cross-validation method. The resulting classification accuracy on both agar cultures was over 99%. The commission errors ranged from 0.11% to 1.19%. The omission errors were between 0% and 1.38%. A future study is needed to develop an imaging system that can deal with more practical culture conditions. In addition, a time-base study for early detection is needed.

Acknowledgements

The authors would like to express a deep appreciation to Johnna Garish for her support and help for this project.

References

Comaniciu, D., V. Ramesh, and P. Meer. 2000. Real-time tracking of non-rigid objects using mean shift. *Proc. IEEE Conf. on Computer Vision and Pattern Recognition* 2: 142-149. Piscataway, N.J.: IEEE.

Landgrebe, D. A. 2003. *Signal Theory Methods in Multispectral Remote Sensing.* New York, N.Y.: Wiley.

Lawrence, K. C., B. Park, W. R. Windham, and C. Mao. 2003. Calibration of a pushbroom hyperspectral imaging system for agricultural inspection. *Trans. ASAE* 46(2): 513-521.

Line, J. E. 2001. Development of a selective differential agar for isolation and enumeration of *Campylobacter* spp. *J. Food Protection* 64(11): 1711-1715.

Mao, C. 2000. Focal plane scanner with reciprocating spatial window. U.S. Patent, 6,166,373.

Mouwen, D. J. M., M. J. B. M. Weijtens, R. Capita, C. Alonso-Calleja, and M. Prieto. 2005. Discrimination of enterobacterial repetitive intergenic consensus PCR types of *Campylobacter coli* and *Campylobacter jejuni* by Fourier transform infrared spectroscopy. *Applied and Environ. Microbiol.* (71)8: 4318-4324.

Naumann, D. 2000. Infrared spectroscopy in microbiology. In *Encyclopedia of Analytical Chemistry*, 102-131. Oxford, U.K.: John Wiley and Sons, Ltd.

Press, W. H., S. A. Teukolsky, W. T. Vettering, and B. P. Flannery. 2002. *Numerical Recipes in C++: The Art of Scientific Computing.* 2nd ed. Cambridge, U.K.: Cambridge University Press.

Siragusa, G. R., J. E. Line, L. L. Brooks, T. Hutchinson, J. D. Laster, and R. O. Apple. 2004. Serological methods and selective agars to enumerate *Campylobacter* from broiler carcasses: Data from inter- and intralaboratory analyses. *J. Food Protection* 67(5): 901-907.

Stern, N. J., J. E. Line, and H. C. Chen. 2001a. *Campylobacter.* In *Compendium of Methods for the Microbiological Examination of Foods.* 3rd ed. Washington, D.C.: American Public Health Association.

Stern, N. J., P. Fedorka-Cray, J. S. Bailey, N. A. Cox, S. E. Craven, K. L. Hiett, M. T. Musgrove, S. Ladely, D. Cosby, and G. C. Mead. 2001b. Distribution of *Campylobacter* spp. in selected U.S. poultry production and processing operations. *J. Food Protection* 64(11): 1705-1710.

USDA-FSIS. 2005. Analytical utility of *Campylobacter* methodologies. Washington, D.C.: USDA-FSIS National Advisory Committee on Microbiological Criteria for Foods. Available at: www.fsis.usda.gov. Accessed 13 December 2007.

Webb, A. 2002. *Statistical Pattern Recognition.* 2nd ed. New York, N.Y.: Wiley.

Yao, H., Z. Hruska, K. DiCrispino, K. Brabham, D. Lewis, and J. Beach. 2005. Differentiation of fungi using hyperspectral imagery for food inspection. ASAE Paper No. 053127. St. Joseph, Mich.: ASABE.